

STIC-Biotech/ChemLib

166919

From: Riggins, Patrick S.
Sent: Tuesday, September 27, 2005 8:58 AM
To: STIC-Biotech/ChemLib
Subject: 10714212

Contacts: STIC

Please search SEQ ID NO: 21 and SEQ ID NO: 22 against the commercial protein databases.

Thank you
Patrick Riggins
Examiner
Art Unit 1633
REM 2D60
(571) 272-6102

RECEIVED
SEP 27 2005
ECH/CHEM. DIVISION
(STIC)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 29, 2005, 19:30:39 ; Search time 118 Seconds
(without alignments)
62.275 Million cell updates/sec

Title: US-10-714-212-21

Perfect score: 91
Sequence: 1 ESEBKTASGIVLPDSAKK 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	19	4	AAU04230
2	88	96.7	19	4	ABBA9470
3	85	93.4	19	4	AAU04229
4	85	93.4	19	4	AAU04231
5	85	93.4	40	2	AAU04231
6	81	89.0	19	4	AAU04238
7	80	87.9	19	4	AAU04238
8	80	87.9	19	4	AAU04239
9	80	87.9	19	4	AAU04239
10	78	85.7	95	8	ADS28875
11	78	85.7	104	8	ADS28888
12	76	83.5	19	4	AAU04219
13	76	83.5	19	4	AAU04226
14	76	83.5	98	8	ADS23298
15	76	83.5	122	8	ADN17513
16	75	82.4	19	4	AAU04233
17	75	82.4	101	2	AAU16679
18	75	82.4	104	8	ADS42281
19	73	80.2	19	4	AAU04240
20	73	80.2	51	2	AAU04240
21	73	80.2	72	2	AAU04240
22	73	80.2	100	8	ADP03527
23	73	80.2	104	5	ABU52127
24	73	80.2	118	2	AAU04238
25	73	80.2	118	2	AAU04238

26	73	80.2	118	2	AAU06731	AAU06731 H. pylori
27	73	80.2	118	4	AAU05688	AAU05688 Helicobac
28	73	80.2	118	4	AAU05872	AAU05872 Helicobac
29	73	80.2	118	7	ADG75218	ADG75218 Helicobac
30	73	80.2	118	7	ADG75220	ADG75220 Helicobac
31	72	79.1	89	6	ADS21261	ADS21261 Bacterial
32	72	79.1	96	6	ABP77243	ABP77243 N. gonorr
33	72	79.1	96	8	ADS20982	ADS20982 Bacterial
34	72	79.1	99	4	AAU04224	AAU04224 C. glutam
35	72	79.1	99	4	AAU04224	AAU04224 C. glutam
36	72	79.1	130	4	AAU05728	AAU05728 Propionib
37	72	79.1	130	6	ABW42247	ABW42247 Propionib
38	71	78.0	19	4	AAU04222	AAU04222 Chaperonin
39	71	78.0	19	4	AAU04224	AAU04224 Chaperonin
40	71	78.0	95	8	ADS43337	ADS43337 Bacterial
41	70	76.9	92	8	ADN20296	ADN20296 Bacterial
42	70	76.9	98	8	ADS25886	ADS25886 Bacterial
43	70	76.9	98	8	ADS26231	ADS26231 Bacterial
44	70	76.9	98	8	ADS22726	ADS22726 Bacterial
45	70	76.9	98	8	ADS25173	ADS25173 Bacterial

ALIGNMENTS

RESULT 1	AAU04230	AAU04230 standard; peptide; 19 AA.
ID	AAU04230	
AC	AAU04230	
DT	23-OCT-2001	(first entry)
DE	Chaperonin binding domain #19.	
XX		
XX	Chaperonin binding domain; therapeutic; protein expression;	
XX	growth factor; cytokine; ligand; receptor; inhibitor; vaccine;	
XX	industrial protein.	
OS	Unidentified.	
XX		
PN	WO200148227-A1.	
XX		
PD	05-JUL-2001.	
XX		
PF	14-DEC-2000; 2000WO-US034055.	
XX		
PR	23-DEC-1999; 99US-00470830.	
PA	(GENV) GENENCOR INT INC.	
PI	Joachim A, Donnelly M;	
XX		
DR	WPI; 2001-425674/45.	
XX		
PT	Producing proteins, particularly toxic proteins in host cells, such as	
PT	bacterial cells, by using a chaperonin binding domain in expression	
PT	systems designed for the production of the proteins.	
XX		
XX	Claim 9; Page 8; 26pp; English.	
CC	The sequence represents the amino acid sequence of chaperonin binding	
CC	domain #19. The sequence is used in a method of producing a protein in a	
CC	host cell, comprising culturing a host cell comprising a nucleic acid	
CC	encoding an isolated chaperonin binding domain (CBD), associated with a	
CC	second nucleic acid encoding the protein and a third nucleic acid	
CC	encoding a chaperonin. This is carried out under conditions suitable for	
CC	expression of the nucleic acids, where CBD is capable of binding to the	
CC	chaperonin. The method is useful for producing a protein, particularly	
CC	toxic protein in a bacterial cell, especially Escherichia coli. CBD is	
CC	useful for producing heterologous proteins, peptides or polypeptides in a	
CC	host cell, including heterototally significant proteins, such as growth	
CC	factors, cytokines, ligands, receptors and inhibitors, vaccines,	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 19:34:24 ; Search time 24.5 Seconds
(without alignments)
74.617 Million cell updates/sec

Title: US-10-714-212-21

Perfect score: 91
Sequence: 1 ESEKTAGIPLPSAXEK 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	94	2	A41884
2	88	96.7	94	2	A41333
3	88	96.7	94	2	AD1704
4	85	93.4	88	2	PC4238
5	85	93.4	94	2	JC1479
6	85	93.4	94	2	A49855
7	85	93.4	94	2	A83720
8	80	87.9	103	1	BVTCGS
9	80	87.9	103	1	A36721
10	76	83.5	98	2	JN0510
11	76	83.5	101	2	JN0513
12	76	83.5	122	2	B70489
13	75	82.4	98	1	A43827
14	75	82.4	98	1	D95311
15	75	82.4	98	2	AP3640
16	75	82.4	104	2	S35310
17	74	81.3	103	2	T06830
18	73	80.2	95	2	B71668
19	73	80.2	106	2	S75563
20	73	80.2	118	2	C71986
21	73	80.2	118	2	S61396
22	72	79.1	95	2	I40330
23	72	79.1	96	2	G81019
24	72	79.1	96	2	G81964
25	72	79.1	103	2	AP2263
26	71	78.0	88	2	A71253
27	71	78.0	95	2	A41872
28	71	78.0	95	2	G97232
29	71	78.0	104	2	S35308

30	71	78.0	120	2	F75499	chaperonin - Deino
31	70	76.9	92	2	G72367	groES protein - Th
32	70	76.9	94	2	S68248	chaperonin groES -
33	70	76.9	98	2	C97442	10K chaperonin (pr
34	70	76.9	98	2	AE2660	10 KD chaperonin (
35	70	76.9	98	2	A69197	heat shock protein
36	69	75.8	100	1	BVMTB	chaperonin groES -
37	69	75.8	100	1	BVMTB	chaperonin groES -
38	69	75.8	102	2	A41325	heat shock protein
39	69	75.8	102	2	S37565	chaperonin groES -
40	68	74.7	95	2	JC2563	heat shock protein
41	68	74.7	99	2	A97821	10K chaperonin (im
42	68	74.7	99	2	C87334	chaperonin, 10 kDa
43	66	72.5	96	2	S70668	chaperonin groES -
44	65	71.4	86	2	F81328	10 KD chaperonin (
45	65	71.4	94	2	JN0600	heat shock protein

ALIGNMENTS

RESULT 1

A41884 heat shock protein (chaperonin) groES - Bacillus subtilis

N/Alternate names: class I heat-shock protein

C/Species: Bacillus subtilis

C/Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 09-Jul-2004

C/Accession: A41884; A41885; JCI371; E69637

R/Li: M.; Wong, S.L. 1981-1992, 1992

J. Bacteriol. 174, 3993-3999, 1992

A>Title: Cloning and characterization of the groESL operon from Bacillus subtilis.

A/Reference number: A41884; MUID:92283753; PMID:1350776

A/Accession: A41884

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-94 <LIL>

A/Cross-references: UNIPROT:P28599; GB:M61132; NID:G143025; PIDN:AAA22502.1; PID:G14302

A/Experimental source: 168, tRPC2

A/Note: this sequence is inconsistent with the nucleotide translation

R/Schmidt, A.; Schieswohl, M.; Volker, U.; Hecker, M.; Schumann, W.

J. Bacteriol. 174, 3993-3999, 1992

A>Title: Cloning, sequencing, mapping, and transcriptional analysis of the groESL opero

A/Reference number: A41885; MUID:92283754; PMID:1350777

A/Accession: A41885

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-94 <SCH>

A/Cross-references: GB:M64965; NID:G143061; PIDN:AAA22530.1; PID:G143062

A/Experimental source: MB1

A/Note: sequence extracted from NCBI backbone (NCBIN:105432, NCBIP:105433)

R/Torawa, Y.; Yoshikawa, H.; Kawamura, F.; Iraya, M.; Takahashi, H.

Biosci. Biotechnol. Biochem. 56, 1995-2002, 1992

A>Title: Isolation and characterization of the groES and groEL genes of Bacillus subtil

A/Reference number: JCI371; MUID:93129852; PMID:1369494

A/Accession: JCI371

A/Molecule type: DNA

A/Residues: 1-94 <TOZ>

A/Cross-references: GB:D10972; GB:D01157; NID:G4433779; PIDN:BA22518.1; PID:G4433780

A/Experimental source: Strain Marburg 168

A/Note: the authors translated the codon ACT for residue 20 as Tyr

R/Kunst, P.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berle

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinis

A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon

A/Authors: Schleich, S.; Schroeder, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tobato, V.; Uchiyama

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 19:33:34 ; Search time 113.5 Seconds
(without alignments)
85.723 Million cell updates/sec

Title: US-10-714-212-21
Sequence: 1 ESEKTPASGIVLPDSAKK 19

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Uniprot 03:.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	91	100.0	94 1 CH10_BACSV	P28599 bacillus su
2	91	100.0	94 2 O65MZ9	O65MZ9 bacillus li
3	88	96.7	94 1 CH10_LISTN	O92909 listeria in
4	88	96.7	94 1 CH10_LISMO	O93907 listeria mo
5	88	96.7	94 2 O71XU5	O71XU5 listeria mo
6	85	93.4	94 1 CH10_BACHD	O50304 bacillus ha
7	85	93.4	94 1 CH10_BACP3	P26210 bacillus ps
8	85	93.4	94 1 CH10_BACST	O07200 bacillus st
9	85	93.4	94 1 CH10_BACR	O8V985 bacillus th
10	85	93.4	94 2 O7DUL2	O7DUL2 bacillus sp
11	85	93.4	94 2 O9EZV5	O9EZV5 bacillus st
12	84	92.3	94 2 O7TU43	O7TU43 prochloroco
13	82	90.1	94 2 O63GV8	O63GV8 bacillus ce
14	82	90.1	94 2 O73ES0	O73ES0 bacillus ce
15	82	90.1	94 2 O81VE2	O81VE2 bacillus an
16	82	90.1	94 2 O6HPC8	O6HPC8 bacillus th
17	82	90.1	96 2 O81AB1	O81AB1 bacillus ce
18	81	89.0	93 1 CH10_OCEIH	O8CX14 oceanobacill
19	81	89.0	103 2 O7TTX2	O7TTX2 synechococc
20	81	89.0	103 2 O7TUS3	O7TUS3 prochloroco
21	81	89.0	103 2 O7TV92	O7TV92 prochloroco
22	81	89.0	103 2 O7TV92	P07889 synechococc
23	80	87.9	103 1 CH10_SYNP6	P22880 synechococc
24	80	87.9	103 1 CH10_SYNP7	O7UN98 rhodospirill
25	78	85.7	104 1 O7UM98	O98110 rhizobium l
26	78	85.7	104 2 O6W164	O6W164 rhizobium s
27	78	85.7	104 2 O6W164	O6W164 rhizobium s
28	76	83.5	89 1 CH10_PORGI	P42376 porphyromon
29	76	83.5	94 1 CH10_CLOBO	O8KJ25 clostridium
30	76	83.5	97 2 O67KB7	O67KB7 symbiodacte
31	76	83.5	98 1 CH11_RHIME	P35473 rhizobium m

ALIGNMENTS

32	76	83.5	98 1	CH13_RHILO	O9848 rhizobium 1
33	76	83.5	98 1	CH14_RHILO	O9843 rhizobium 1
34	76	83.5	98 1	CH15_RHILO	O9810 rhizobium 1
35	76	83.5	98 2	O799Q4	O799Q4 rhizobium 1
36	76	83.5	98 2	O6FYU9	O6FYU9 bartonella
37	76	83.5	98 2	O6H8L3	O6H8L3 bartonella
38	76	83.5	103 1	CH10_SYNP8	O81373 synechococc
39	76	83.5	104 1	CH15_RHIME	P35474 rhizobium m
40	76	83.5	122 1	CH10_AQUAE	O67942 aquifex aeo
41	75	82.4	93 2	O7TTZ0	O7TTZ0 rhodospirill
42	75	82.4	95 2	O7TAL5	O7TAL5 desulfotolylr
43	75	82.4	95 2	O6ARV5	O6ARV5 desulfotolylr
44	75	82.4	98 1	CH10_BRUME	P25968 bruceella me
45	75	82.4	98 1	CH14_RHIME	O92423 rhizobium m

RESULT 1
CH10_BACSV STANDARD; PRT; 94 AA.
ID CH10_BACSV
AC P28599; O08341;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-JUL-1993 (Rel. 45, Last annotation update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 10 kDa chaperonin (protein Cpn10) (groES protein).
GN Name=groS; Synonyms=groES; mopB; OrderedLocNames=BSU06020;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92283754; PubMed=1350777;
RA Schmidt A., Schlesswohl M., Voelker U., Hecker M., Schumann W.;
RT "Cloning, sequencing, mapping, and transcriptional analysis of the
RT groES operon from Bacillus subtilis.";
RL J. Bacteriol. 174:3993-3999(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RA MEDLINE=92283753; PubMed=1350776;
RT Li M., Wong S.L.;
RT "Cloning and characterization of the groESL operon from Bacillus
RT subtilis.";
RL J. Bacteriol. 174:3981-3992(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=168 / Marburg;
RA MEDLINE=93129852; PubMed=1369494;
RT Tozawa Y., Yoshikawa H., Kawamura F., Itaya M., Takahashi H.;
RT "Isolation and characterization of the groE and groEL genes of
RT Bacillus subtilis Marburg.";
RL Biosci. Biotechnol. Biochem. 56:1995-2002(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=168 / JH642;
RA MEDLINE=97346038; PubMed=9202461;
RT Sadaie Y., Yata K., Fujita M., Sagai H., Itaya M., Kasahara Y.,
RA Ogasawara N.;
RT "Nucleotide sequence and analysis of the phoB-rrnE-groESL region of
RT the Bacillus subtilis chromosome.";
RL Microbiology 143:1861-1866(1997).
RN [5]
RP SEQUENCE OF 5-94 FROM N.A.
RX STRAIN=168 / Marburg;
RA MEDLINE=98116660; PubMed=9455482;
RT Kasahara Y., Nakai S., Ogasawara N., Yata K., Sadaie Y.;
RT "Sequence analysis of the groESL-cotA region of the Bacillus subtilis
RT genome, containing the restriction/modification system genes.";
RL DNA Res. 4:335-339(1997).
RN [6]
RP SEQUENCE FROM N.A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW Model

Run on: September 29, 2005, 19:35:05 / Search time 30.5 Seconds
(without alignments)
46.503 Million cell updates/sec

Title: US-10-714-212-21

Perfect score: 91

Sequence: 1 ESEKTAGSIVLPDSAKK 19

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiller.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	19	4	US-09-470-830A-21
2	85	93.4	19	4	US-09-470-830A-20
3	85	93.4	19	4	US-09-470-830A-22
4	80	87.9	19	4	US-09-470-830A-29
5	80	87.9	19	4	US-09-470-830A-30
6	80	87.9	19	4	US-08-467-822-38
7	80	87.9	19	4	US-08-432-697-38
8	80	87.9	19	4	US-08-466-248-38
9	76	83.5	19	4	US-09-470-830A-10
10	76	83.5	19	4	US-09-470-830A-17
11	75	82.4	19	4	US-09-470-830A-12
12	75	82.4	19	4	US-09-470-830A-24
13	73	80.2	19	4	US-09-470-830A-31
14	73	80.2	19	4	US-08-467-822-29
15	73	80.2	19	4	US-08-432-697-29
16	73	80.2	19	3	US-08-466-248-29
17	72	79.1	19	4	US-09-603-208A-4
18	72	79.1	19	4	US-09-470-830A-13
19	71	78.0	19	4	US-09-470-830A-15
20	70	76.9	19	4	US-09-470-830A-14
21	69	75.8	19	4	US-09-470-830A-19
22	69	75.8	19	4	US-09-470-830A-23
23	69	75.8	19	4	US-09-470-830A-26
24	69	75.8	19	4	US-09-613-303-35
25	69	75.8	19	4	US-10-267-311-35
26	65	71.4	19	4	US-09-470-830A-25
27	65	71.4	11	3	US-09-134-001C-4686

28	64	70.3	19	4	US-09-470-830A-6	Sequence 6, Appl
29	64	70.3	19	4	US-09-470-830A-9	Sequence 9, Appl
30	64	70.3	19	4	US-09-470-830A-11	Sequence 11, Appl
31	64	70.3	94	2	US-08-467-822-37	Sequence 37, Appl
32	64	70.3	94	3	US-08-432-697-37	Sequence 37, Appl
33	64	70.3	94	3	US-08-466-248-37	Sequence 37, Appl
34	64	70.3	140	4	US-09-540-236-2611	Sequence 2611, Ap
35	64	70.3	145	4	US-09-252-991A-24284	Sequence 24284, A
36	61	67.0	106	4	US-09-328-352-5510	Sequence 5510, Ap
37	60	65.9	19	4	US-09-470-830A-8	Sequence 8, Appl
38	60	65.9	19	4	US-09-470-830A-27	Sequence 27, Appl
39	59	64.8	19	4	US-09-470-830A-7	Sequence 7, Appl
40	59	64.8	93	2	US-08-467-822-36	Sequence 36, Appl
41	59	64.8	93	3	US-08-432-697-36	Sequence 36, Appl
42	59	64.8	93	3	US-08-466-248-36	Sequence 36, Appl
43	58.5	64.3	102	4	US-09-438-185A-137	Sequence 137, App
44	58	63.7	119	4	US-09-107-532A-6969	Sequence 6969, Ap
45	57.5	63.2	102	4	US-09-198-452A-154	Sequence 154, App

ALIGNMENTS

```

RESULT 1
US-09-470-830A-21
; Sequence 21, Application US/09470830A
; Patent No. 6677139
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; FILE REFERENCE: GC559-PCT
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/470, 830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-470-830A-21

Query Match      100.0%; Score 91; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ESEKTAGSIVLPDSAKK 19
        |||||
Db       1 ESEKTAGSIVLPDSAKK 19

RESULT 2
US-09-470-830A-20
; Sequence 20, Application US/09470830A
; Patent No. 6677139
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; FILE REFERENCE: GC559-PCT
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/470, 830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Thermocactinomyces sp.
US-09-470-830A-20

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 19:38:05; Search time 117 Seconds
(without alignments) 67.412 Million cell updates/sec

Title: US-10-714-212-21

Perfect score: 91

Sequence: 1 ESEKTAGIVLPDSAKK 19

Scoring table: BLOSUM62

Searched: Gapop 10.0, Gapext 0.5

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*

12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*

18: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*

19: /cgn2_6/ptodata/1/pubppa/US11_PUBCOMB.pep:*

20: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep:*

21: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*

22: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	19	US-10-714-212-21	Sequence 21, Appl
2	85	93.4	19	US-10-714-212-20	Sequence 20, Appl
3	85	93.4	19	US-10-714-212-22	Sequence 22, Appl
4	81	89.0	94	US-10-369-493-10863	Sequence 10863, A
5	80	87.9	19	US-10-714-212-29	Sequence 29, Appl
6	80	87.9	19	US-10-714-212-30	Sequence 30, Appl
7	78	85.7	95	US-10-369-493-117908	Sequence 117908, A
8	78	85.7	104	US-10-369-493-11921	Sequence 11921, A
9	76	83.5	19	US-10-714-212-10	Sequence 10, Appl
10	76	83.5	19	US-10-714-212-17	Sequence 17, Appl
11	76	83.5	98	US-10-369-493-12331	Sequence 12331, A

12	76	83.5	122	15	US-10-369-493-166	Sequence 166, App
13	75	82.4	19	15	US-10-714-212-12	Sequence 12, Appl
14	75	82.4	19	15	US-10-714-212-24	Sequence 24, Appl
15	75	82.4	104	15	US-10-369-493-20711	Sequence 20711, A
16	73	80.2	19	15	US-10-714-212-31	Sequence 31, Appl
17	73	80.2	118	9	US-09-815-242-11281	Sequence 11281, A
18	73	80.2	118	9	US-09-815-242-11465	Sequence 11465, A
19	73	80.2	118	18	US-10-953-901-419	Sequence 419, App
20	73	80.2	118	18	US-10-953-901-421	Sequence 421, App
21	72	79.1	89	15	US-10-369-493-10294	Sequence 10294, A
22	72	79.1	96	15	US-10-369-493-10015	Sequence 10015, A
23	72	79.1	99	9	US-09-738-626-4164	Sequence 4164, App
24	71	78.0	19	15	US-10-714-212-13	Sequence 13, Appl
25	71	78.0	19	15	US-10-714-212-15	Sequence 15, Appl
26	71	78.0	95	15	US-10-369-493-21767	Sequence 21767, A
27	70	76.9	92	15	US-10-714-212-14	Sequence 14, Appl
28	70	76.9	92	15	US-10-369-493-2949	Sequence 2949, App
29	70	76.9	98	15	US-10-369-493-11759	Sequence 11759, A
30	70	76.9	98	15	US-10-369-493-14206	Sequence 14206, A
31	70	76.9	98	15	US-10-369-493-14459	Sequence 14459, A
32	70	76.9	98	15	US-10-369-493-14919	Sequence 14919, A
33	70	76.9	98	15	US-10-369-493-15264	Sequence 15264, A
34	69	75.8	19	15	US-10-714-212-19	Sequence 19, Appl
35	69	75.8	19	15	US-10-714-212-23	Sequence 23, Appl
36	69	75.8	19	15	US-10-714-212-26	Sequence 26, Appl
37	69	75.8	100	9	US-09-712-363-272	Sequence 272, App
38	69	75.8	102	16	US-10-332-512A-11	Sequence 11, Appl
39	69	75.8	102	14	US-10-156-761-12525	Sequence 12525, A
40	69	75.8	102	14	US-10-156-761-12525	Sequence 12525, A
41	69	75.8	198	14	US-10-267-311-35	Sequence 35, Appl
42	68	74.7	95	17	US-10-679-956-35	Sequence 35, Appl
43	67	73.6	101	15	US-10-369-493-7916	Sequence 7916, App
44	66	72.5	96	15	US-10-369-493-20609	Sequence 20609, A
45	66	72.5	97	15	US-10-369-493-16818	Sequence 16818, A
					US-10-369-493-14074	Sequence 14074, A

ALIGNMENTS

RESULT 1

US-10-714-212-21

; Sequence 21, Application US/10714212

; Publication No. US2004007038A1

GENERAL INFORMATION:

APPLICANT: Genencor International, Inc.

TITLE OF INVENTION: Methods for Production of Proteins in

FILE REFERENCE: GCS59-PCT

CURRENT APPLICATION NUMBER: US/10/714,212

PRIOR FILING DATE: 2003-11-13

PRIOR APPLICATION NUMBER: US 09/470,830

NUMBER OF SEQ ID NOS: 44

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 21

LENGTH: 19

TYPE: PRT

ORGANISM: Bacillus subtilis

US-10-714-212-21

Query Match 100.0%; Score 91; DB 15; Length 19;

Best Local Similarity 100.0%; Pred. No. 8.7e-09;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ESEKTAGIVLPDSAKK 19

Db 1 ESEKTAGIVLPDSAKK 19

RESULT 2

US-10-714-212-20

; Sequence 20, Application US/10714212

; Publication No. US2004007038A1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 19:30:39 ; Search time 118 Seconds
(without alignments)
62.275 Million cell updates/sec

Title: US-10-714-212-22

Perfect score: 93

Sequence: 1 ETEBKTAGSIVLPDTAKK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_16Dec04:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003as:*
8: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	100.0	19	AAU04229	Chaperonin
2	93	100.0	19	AAU04231	Chaperonin
3	93	100.0	40	AAU04231	Heat shock
4	88	94.6	94	AAU04230	GroES-like
5	85	91.4	19	AAU04230	Chaperonin
6	84	90.3	94	AAU04230	Chaperonin
7	81	87.1	95	AAU04230	Chaperonin
8	81	87.1	104	AAU04230	Chaperonin
9	79	84.9	122	AAU04230	Chaperonin
10	78	83.9	19	AAU04230	Chaperonin
11	78	83.9	19	AAU04230	Chaperonin
12	78	83.9	19	AAU04230	Chaperonin
13	78	83.9	104	AAU04230	Chaperonin
14	77	82.8	19	AAU04230	Chaperonin
15	77	82.8	19	AAU04230	Chaperonin
16	77	82.8	94	AAU04230	Chaperonin
17	77	82.8	94	AAU04230	Chaperonin
18	76	81.7	101	AAU04230	Chaperonin
19	75	80.6	89	AAU04230	Chaperonin
20	75	80.6	130	AAU04230	Chaperonin
21	75	80.6	130	AAU04230	Chaperonin
22	74	79.6	19	AAU04230	Chaperonin
23	74	79.6	100	AAU04230	Chaperonin
24	73	78.5	96	AAU04230	Chaperonin
25	72	77.4	19	AAU04230	Chaperonin

26	72	77.4	19	AAU04228	Chaperonin
27	72	77.4	95	AAU04228	Bacterial
28	72	77.4	99	AAU05700	MPB-57 pr
29	72	77.4	100	AAU05700	Mycobacte
30	72	77.4	100	AAU05700	Mycobacte
31	72	77.4	100	AAU05700	Mycobacte
32	72	77.4	100	AAU05700	Mycobacte
33	72	77.4	102	AAU05700	Mycobacte
34	72	77.4	102	AAU05700	Mycobacte
35	72	77.4	102	AAU05700	Mycobacte
36	72	77.4	102	AAU05700	Mycobacte
37	72	77.4	102	AAU05700	Mycobacte
38	72	77.4	102	AAU05700	Mycobacte
39	72	77.4	102	AAU05700	Mycobacte
40	72	77.4	102	AAU05700	Mycobacte
41	72	77.4	102	AAU05700	Mycobacte
42	72	77.4	102	AAU05700	Mycobacte
43	72	77.4	102	AAU05700	Mycobacte
44	72	77.4	102	AAU05700	Mycobacte
45	72	77.4	102	AAU05700	Mycobacte

ALIGNMENTS

RESULT 1	AAU04229	standard; peptide: 19 AA.
ID	AAU04229	
AC	AAU04229;	
XX		
DT	23-OCT-2001 (first entry)	
XX		
DE	Chaperonin binding domain #18.	
XX		
KW	Chaperonin binding domain; therapeutic; protein expression; growth factor; cytokine; ligand; receptor; inhibitor; vaccine; industrial protein.	
KW		
OS	Unidentified.	
XX		
PN	WO200148227-A1.	
XX		
PD	05-JUN-2001.	
XX		
PF	14-DEC-2000; 2000WO-US034055.	
XX		
PR	23-DEC-1999; 99US-00470830.	
XX		
PA	(GENV) GENENOR INT INC.	
XX		
PI	Joachim A. Donnelly M;	
XX		
DR	WPI; 2001-425674/45.	
XX		
PT	Producing proteins, particularly toxic proteins in host cells, such as bacterial cells, by using a chaperonin binding domain in expression systems designed for the production of the proteins.	
PT		
PS	Claim 9; Page 8; 26pp; English.	
XX		
CC	The sequence represents the amino acid sequence of chaperonin binding domain #18. The sequence is used in a method of producing a protein in a host cell, comprising culturing a host cell comprising a nucleic acid encoding an isolated chaperonin binding domain (CBD), associated with a second nucleic acid encoding the protein and a third nucleic acid encoding a chaperonin. This is carried out under conditions suitable for expression of the nucleic acids, where CBD is capable of binding to the chaperonin. The method is useful for producing a protein, particularly toxic protein in a bacterial cell, especially Escherichia coli. CBD is useful for producing heterologous proteins, peptides or polypeptides in a host cell, including therapeutically significant proteins, such as growth factors, cytokines, ligands, receptors and inhibitors, vaccines,	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 29, 2005, 19:34:24 ; Search time 24.5 seconds
(without alignments)

74.617 Million cell updates/sec

Title: US-10-714-212-22

Perfect score: 93
Sequence: 1 ETEKTAGSIVLPDTAKEK 19

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	93	100.0	88 2 PC4238	heat shock protein
2	93	100.0	94 2 TC1479	heat shock protein
3	93	100.0	94 2 A49855	heat shock protein
4	93	100.0	94 2 A83720	class I heat-shock
5	85	91.4	94 2 A41884	heat shock protein
6	84	90.3	94 2 A41333	class I heat-shock
7	84	90.3	94 2 AD1704	class I heat-shock
8	79	84.9	122 2 B70489	GROES - Agulifex ae
9	78	83.9	103 1 BVCYCS	chaperonin groES -
10	78	83.9	103 1 A36721	groES protein - Sy
11	78	83.9	104 2 S35310	heat shock cognate
12	77	82.8	98 2 JN0510	heat shock protein
13	77	82.8	101 2 JN0513	heat shock protein
14	76	81.7	95 2 B71658	10 kd chaperonin (
15	76	81.7	98 1 A43827	chaperonin groES -
16	76	81.7	98 1 D95311	groES2 chaperonin
17	76	80.6	98 2 AFG640	10K chaperonin gro
18	75	80.6	103 2 T06830	chaperonin groES -
19	74	79.6	104 2 S35308	heat shock protein
20	73	78.5	103 2 AR2263	chaperonin groES (
21	72	77.4	88 2 A71253	probable chaperon
22	72	77.4	100 1 BVMY7B	chaperonin groES -
23	72	77.4	100 1 BVMY7B	chaperonin groES -
24	72	77.4	102 2 A41325	heat shock protein
25	72	77.4	102 2 S37565	chaperonin groES -
26	71	76.3	94 2 S68248	chaperonin groES -
27	71	76.3	95 2 UC2563	heat shock protein
28	71	76.3	98 2 C97442	10K chaperonin (pr
29	71	76.3	98 2 AE2660	10 KD chaperonin (

ALIGNMENTS

RESULT 1

PC4238
heat shock protein GROES - Bacillus sp. (fragment)

N/Alternate names: molecular chaperone 60 GROES

C/Species: Bacillus sp.

C/Date: 20-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004

C/Accession: PC4238; PC6023

R/Xu, Y.; Kobayashi, T.; Kudo, T.

BioSci. Biotechnol. Biochem. 60, 1633-1636, 1996

A/Title: Molecular cloning and nucleotide sequence of the groEL gene from the alkaliph

A/Reference number: JCS130; MUID:97141316; PMID:8987660

A/Accession: PC4238

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-88 <XU>

A/Cross-references: UNIPROT:O50304; DDBJ:D55630; NID:G1682949; PIDN:BA09493.1; PID:G1

A/Experimental source: strain C-125

R/Xu, Y.; Zhou, P.J.

Acta Microbiol. Sin. 36, 241-249, 1996

A/Title: Phylogeny of molecular chaperone 60 proteins.

A/Reference number: JC6063

A/Accession: PC6023

A/Molecule type: DNA

A/Experimental source: C-125 strain BD224

C/Comment: This protein plays a role in protein folding by binding to the complex of a

A/Gene: groES

C/Superfamily: chaperonin groES

C/Keywords: molecular chaperone

Query Match 100.0%; Score 93; DB 2; Length 88;

Best Local Similarity 100.0%; Pred. No. 2,5e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 ETEKTAGSIVLPDTAKEK 19

9 ETEKTAGSIVLPDTAKEK 27

RESULT 2

JC1479

heat shock protein TGROES - thermophilic bacterium PS-3

N/Alternate names: heat shock 12K protein; hsp10

C/Species: Thermophilic bacterium PS-3

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 12-Sep-1997

C/Accession: JC1479; JQ1194

R/Ohta, T.; Honda, K.; Saito, K.; Hayashi, H.; Tano, H.; Hamamoto, T.; Kagawa, Y.

Biochem. Biophys. Res. Commun. 191, 550-557, 1993

A/Title: Heat shock promoter of thermophilic chaperonin operon.

A/Reference number: JC1479; MUID:93213292; PMID:8096382

A/Accession: JC1479

30	71	76.3	98 2 A36917	heat shock protein
31	71	76.3	99 2 A97821	10K chaperonin (im
32	71	76.3	106 2 S77563	chaperonin groES -
33	71	76.3	118 2 C71986	10Kda chaperone -
34	71	76.3	118 2 S61396	chaperonin groES -
35	70	75.3	96 2 G81019	chaperonin 10 kda
36	70	75.3	96 2 G81964	chaperonin 10 kda
37	69	74.2	96 2 C87334	chaperonin, 10 kda
38	69	74.2	99 2 S70668	chaperonin groES -
39	68	73.1	95 2 I40330	Cpn10 protein (Gro
40	67	72.0	95 2 A41872	heat shock protein
41	67	72.0	95 2 G97232	co-chaperonin GroE
42	67	72.0	120 2 F75499	chaperonin - Deino
43	66	71.0	92 2 G72367	chaperonin - Th
44	63	67.7	86 2 F81328	10 kd chaperonin (
45	63	67.7	253 2 T52613	chaperonin 21 prec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2005, 19:33:34 ; Search time 113.5 Seconds
(without alignments)
85.723 Million cell updates/sec

Title: US-10-714-212-22

Perfect score: 93
Sequence: 1 ETEKTAGSIVLPDPAK 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	93	100.0	94	CH10_BACHD	O50304 bacillus ha
2	93	100.0	94	CH10_BACR3	P26210 bacillus ps
3	93	100.0	94	CH10_BACR1	O07200 bacillus st
4	93	100.0	94	CH10_BACR2	O8V985 bacillus st
5	93	100.0	94	CH10_BACR2	O7D312 bacillus sp
6	93	100.0	94	O9EZV5	O9EZV5 bacillus st
7	85	91.4	94	CH10_BACSU	P28599 bacillus su
8	85	91.4	94	O63GV8	O63GV8 bacillus su
9	85	91.4	94	O65M29	O65M29 bacillus li
10	85	91.4	94	O73E80	O73E80 bacillus li
11	85	91.4	94	O81VE2	O81VE2 bacillus an
12	85	91.4	94	O6HPC8	O6HPC8 bacillus th
13	85	91.4	96	O814B1	O814B1 bacillus ce
14	84	90.3	94	CH10_LISIN	O929U9 listeria in
15	84	90.3	94	CH10_LISMO	O929U9 listeria mo
16	84	90.3	94	O71XU5	O71XU5 listeria mo
17	82	88.2	103	O7TUX2	O7TUX2 synecococc
18	82	88.2	103	O7TUX3	O7TUX3 synecococc
19	82	88.2	103	O7TV92	O7TV92 synecococc
20	81	87.1	104	CH12_RHIL0	O98110 rhizobium l
21	81	87.1	104	O6W164	O6W164 rhizobium s
22	81	87.1	104	O6W1D4	O6W1D4 rhizobium s
23	79	84.9	97	O67KB7	O67KB7 symbiobacte
24	79	84.9	122	CH10_AQUAE	O67942 aquifex aeo
25	78	83.9	103	CH10_STNPE	P07889 synecococc
26	78	83.9	103	CH10_STNPE	P22880 synecococc
27	78	83.9	103	O7TU43	O7TU43 synecococc
28	78	83.9	104	CH11_BRAJA	P77828 bradyrhizob
29	78	83.9	104	CH12_RHOPA	P80367 rhodopseud
30	78	83.9	104	CH13_BRAJA	P35864 bradyrhizob
31	78	83.9	104	O89IK9	O89IK9 bradyrhizob

ALIGNMENTS

RESULT 1	CH10_BACHD	STANDARD;	PRT;	94 AA.
AC	O50304;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	10 kDa chaperonin (protein Cpn10) (groES protein).			
GN	Name=groS; Synonyms=groES, mopB; OrderedLocustNames=BH0561;			
OS	Bacillus halodurans.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=86665;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C-125 / JCM 9153;			
RX	MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;			
RA	Xu Y., Kobayashi T., Kudo T.;			
RA	Takemi H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,			
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,			
RA	Horikoshi K.;			
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus			
RT	halodurans and genomic sequence comparison with Bacillus subtilis.";			
RL	Nucleic Acids Res. 28:4317-4331(2000).			
RN	[2]			
RP	SEQUENCE OF 7-94 FROM N.A.			
RC	STRAIN=C-125 / JCM 9153;			
RX	MEDLINE=9714316; PubMed=8987660;			
RA	Xu Y., Kobayashi T., Kudo T.;			
RA	"Molecular cloning and nucleotide sequence of the groEL gene from the			
RT	alkaliphilic Bacillus sp. strain C-125 and reactivation of thermally			
RT	inactivated alpha-glucosidase by recombinant GroEL.";			
RL	Biotechol. Biochem. 60:1633-1636(1996).			
CC	- FUNCTION: Binds to Cpn60 in the presence of Mg-ATP and suppresses			
CC	the ATPase activity of the latter.			
CC	- SUBUNIT: Heptamer of 7 subunits arranged in a ring (by			
CC	similarity).			
CC	- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).			
CC	- INDUCTION: By heat shock.			
CC	- SIMILARITY: Belongs to the groES chaperonin family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; AP001508; BAB04280.1;			O8B96 heliobacilli
DR	EMBL; D55630; BAA09493.1;			O8C14 oceanobacilli
DR	PIR; A83720; A83720.			P35473 rhizobium m
DR	PIR; PC4238; PC4238.			O98A8 rhizobium l
DR	HSSP; P09621; IHXS.			O98B3 rhizobium l
DR	HAMP; MF 00580; -1.			O9815 rhizobium l
DR	InterPro; IPR001476; Chaperin_Cpn10.			O799Q4 rhizobium l
				O6FY9 bartonella
				O6Y813 bartonella
				O81373 synecococc
				P35474 rhizobium m
				P80469 rickettsia
				O9ZCT6 rickettsia
				O72AL5 desulfovibr

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - protein search, using BW model

Run on: September 29, 2005, 19:35:05 ; Search time 30.5 Seconds
(without alignments)
46,503 Million cell updates/sec

Title: US-10-714-212-22

Perfect score: 93

Sequence: 1 ETEKTAGIVLPPTAKEX 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents, AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	19	4	US-09-470-830A-20
2	93	100.0	19	4	US-09-470-830A-22
3	88	94.6	94	2	US-08-467-822-38
4	88	94.6	94	3	US-08-432-697-38
5	88	94.6	94	3	US-08-466-248-38
6	85	91.4	19	4	US-09-470-830A-21
7	78	83.9	19	4	US-09-470-830A-24
8	78	83.9	19	4	US-09-470-830A-29
9	78	83.9	19	4	US-09-470-830A-30
10	77	82.8	19	4	US-09-470-830A-10
11	77	82.8	19	4	US-09-470-830A-17
12	76	81.7	19	4	US-09-470-830A-12
13	74	79.6	19	4	US-09-470-830A-13
14	72	77.4	19	4	US-09-470-830A-19
15	72	77.4	19	4	US-09-470-830A-23
16	72	77.4	19	4	US-09-470-830A-26
17	72	77.4	198	4	US-09-613-303-35
18	72	77.4	198	4	US-10-267-311-35
19	71	76.3	19	4	US-09-470-830A-14
20	71	76.3	19	4	US-09-470-830A-31
21	71	76.3	118	2	US-08-467-822-29
22	71	76.3	118	3	US-08-432-697-29
23	71	76.3	118	3	US-08-466-248-29
24	68	73.1	99	4	US-09-603-208A-4
25	67	72.0	19	4	US-09-470-830A-15
26	61.5	66.1	102	4	US-09-438-185A-137
27	61	65.6	19	4	US-09-470-830A-25

28	61	65.6	111	3	US-09-134-001C-4686	Sequence 4686, Ap
29	60.5	65.1	102	4	US-09-198-452A-154	Sequence 154, App
30	60	64.5	19	4	US-09-470-830A-6	Sequence 6, Appl
31	60	64.5	19	4	US-09-470-830A-9	Sequence 9, Appl
32	60	64.5	19	4	US-09-470-830A-11	Sequence 11, Appl
33	60	64.5	94	2	US-08-467-822-37	Sequence 37, Appl
34	60	64.5	94	3	US-08-432-697-37	Sequence 37, Appl
35	60	64.5	94	3	US-08-466-248-37	Sequence 37, Appl
36	60	64.5	96	4	US-09-902-540-10579	Sequence 10579, A
37	60	64.5	140	4	US-09-540-236-2611	Sequence 2611, Ap
38	60	64.5	145	4	US-09-252-991A-24284	Sequence 24284, A
39	58	62.4	19	4	US-09-470-830A-27	Sequence 27, Appl
40	57.5	61.8	19	4	US-09-470-830A-32	Sequence 32, Appl
41	57	61.3	93	2	US-08-467-822-36	Sequence 36, Appl
42	57	61.3	93	3	US-08-432-697-36	Sequence 36, Appl
43	57	61.3	93	3	US-08-466-248-36	Sequence 36, Appl
44	57	61.3	106	4	US-09-328-352-5510	Sequence 5510, Ap
45	56	60.2	19	4	US-09-470-830A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-470-830A-20
; Sequence 20, Application US/09470830A
; Patent No. 6677139
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; FILE REFERENCE: GC559-PTCT
; CURRENT APPLICATION NUMBER: US/09/470, 830A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/470, 830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Thermococcus sp.
US-09-470-830A-20

Query Match 100.0%; Score 93; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ETEKTAGIVLPPTAKEX 19

Db 1 ETEKTAGIVLPPTAKEX 19

RESULT 2
US-09-470-830A-22
; Sequence 22, Application US/09470830A
; Patent No. 6677139
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; FILE REFERENCE: GC559-PTCT
; CURRENT APPLICATION NUMBER: US/09/470, 830A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/470, 830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-09-470-830A-22

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using BW model

Run on: September 29, 2005, 19:38:05 ; Search time 117 Seconds
(without alignments)
67.412 Million cell updates/sec

Title: US-10-714-212-22

Sequence: 93
1 ETEKTSAGIVLPPTAKEK 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1846076 seqs, 415116000 residues

Total number of hits satisfying chosen parameters: 1846076

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCTI_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	93	100.0	19	US-10-714-212-20	Sequence 20, Appl
2	93	100.0	19	US-10-714-212-22	Sequence 22, Appl
3	85	91.4	19	US-10-714-212-21	Sequence 21, Appl
4	81	87.1	95	US-10-369-493-17908	Sequence 17908, A
5	81	87.1	104	US-10-369-493-11921	Sequence 11921, A
6	79	84.9	122	US-10-369-493-166	Sequence 166, App
7	78	83.9	19	US-10-714-212-24	Sequence 24, Appl
8	78	83.9	19	US-10-714-212-29	Sequence 29, Appl
9	78	83.9	19	US-10-714-212-30	Sequence 30, Appl
10	78	83.9	104	US-10-369-493-20711	Sequence 20711, A
11	77	82.8	19	US-10-714-212-10	Sequence 10, Appl

12	77	82.8	19	US-10-714-212-17	Sequence 17, Appl
13	77	82.8	94	US-10-369-493-10863	Sequence 10863, A
14	77	82.8	98	US-10-369-493-12331	Sequence 12331, A
15	76	81.7	19	US-10-714-212-12	Sequence 12, Appl
16	75	80.6	89	US-10-369-493-10294	Sequence 10294, A
17	74	79.6	19	US-10-714-212-13	Sequence 13, Appl
18	73	78.5	96	US-10-369-493-10015	Sequence 10015, A
19	72	77.4	19	US-10-714-212-19	Sequence 19, Appl
20	72	77.4	19	US-10-714-212-23	Sequence 23, Appl
21	72	77.4	19	US-10-714-212-26	Sequence 26, Appl
22	72	77.4	95	US-10-369-493-21767	Sequence 21767, A
23	72	77.4	100	US-09-712-363-272	Sequence 272, App
24	72	77.4	100	US-10-332-512A-11	Sequence 11, Appl
25	72	77.4	102	US-10-156-761-12525	Sequence 12525, A
26	72	77.4	198	US-10-267-311-35	Sequence 35, Appl
27	72	77.4	198	US-10-679-956-35	Sequence 35, Appl
28	71	76.3	19	US-10-714-212-14	Sequence 14, Appl
29	71	76.3	19	US-10-714-212-31	Sequence 31, Appl
30	71	76.3	98	US-10-369-493-11759	Sequence 11759, A
31	71	76.3	98	US-10-369-493-14206	Sequence 14206, A
32	71	76.3	98	US-10-369-493-14459	Sequence 14459, A
33	71	76.3	98	US-10-369-493-14919	Sequence 14919, A
34	71	76.3	98	US-10-369-493-15264	Sequence 15264, A
35	71	76.3	118	US-09-815-242-11281	Sequence 11281, A
36	71	76.3	9	US-09-815-242-11465	Sequence 11465, A
37	71	76.3	118	US-10-953-901-419	Sequence 419, App
38	71	76.3	118	US-10-953-901-421	Sequence 421, App
39	69	74.2	95	US-10-369-493-7916	Sequence 7916, App
40	69	74.2	96	US-10-369-493-16818	Sequence 16818, A
41	68	73.1	99	US-09-738-626-4164	Sequence 4164, App
42	67	72.0	19	US-10-714-212-15	Sequence 15, Appl
43	67	72.0	94	US-10-369-493-15645	Sequence 15645, A
44	67	72.0	94	US-10-369-493-16020	Sequence 16020, A
45	67	72.0	15	US-10-369-493-16372	Sequence 16372, A

ALIGNMENTS

RESULT 1
US-10-714-212-20
; Sequence 20, Application US/10714212
; Publication No. US20040077038A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Methods for Production of Proteins in
; FILE REFERENCE: GC559-PCT
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 09/470,830
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Thermoactinomyces sp.
US-10-714-212-20

Query Match 100.0%; Score 93; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETEKTSAGIVLPPTAKEK 19
DB 1 ETEKTSAGIVLPPTAKEK 19

RESULT 2
US-10-714-212-22
; Sequence 22, Application US/10714212
; Publication No. US20040077038A1